The Importance of Heterogeneity to the Epidemiology of Tuberculosis

James M Trauer[1], Peter J Dodd[2], M Gabriela M Gomes[3,4], Gabriela B Gomez[5], Rein MGJ Houben[6,7], Emma S McBryde[8], Yayehirad A Melsew[1], Nicolas A Menzies[9], Nimalan Arinaminpathy[10], Sourya Shrestha[11], David W Dowdy[11]

1. School of Public Health and Preventive Medicine, Monash University, Australia
2. Health Economic and Decision Science, the University of Sheffield, UK
3. Liverpool School of Tropical Medicine, United Kingdom
4. CIBIO-InBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Portugal
5. Department of Global Health and Development, London School of Hygiene and Tropical Medicine, UK
6. TB Centre, London School of Hygiene and Tropical Medicine, UK
7. Infectious Disease Epidemiology Department, London School of Hygiene and Tropical Medicine, UK
8. Australian Institute of Tropical Health and Medicine, James Cook University, Townsville, Australia
9. Department of Global Health and Population, Harvard T. H. Chan School of Public Health, USA
10. Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, UK
11. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Corresponding author contact details

James Trauer

james.trauer@monash.edu

+613 99030798

School of Public Health and Preventive Medicine

Monash University

553 St Kilda Road, Melbourne

Australia

Summary

- Heterogeneity in TB burden is driven by the organism, host, environment and distal determinants.
- More reliable data are needed, given inconsistent case ascertainment.
- Targeting high-risk groups is an important consideration in designing interventions, but raises equity and efficiency issues.

Abstract

Although less well-recognised than for other infectious diseases, heterogeneity is a defining feature of TB epidemiology. To advance toward TB elimination, this heterogeneity must be better understood and addressed. Drivers of heterogeneity in TB epidemiology act at the level of the infectious host, organism, susceptible host, environment and distal determinants. These effects may be amplified by social mixing patterns, while the variable latent period between infection and disease may mask heterogeneity in transmission. Reliance on notified cases may lead to misidentification of the most affected groups, as case detection is often poorest where prevalence is highest. Assuming average rates apply across diverse groups and ignoring the effects of cohort selection may result in misunderstanding of the epidemic and the anticipated effects of control measures. Given this substantial heterogeneity, interventions targeting high-risk groups based on location, social determinants or comorbidities could improve efficiency, but raise ethical and equity considerations.

Key words

tuberculosis, heterogeneity, epidemiology, case detection, interventions

Introduction

Although estimates of the global burden of tuberculosis (TB) suggest gradual decline, this aggregate profile masks a patchy, heterogeneous epidemic that predominantly afflicts society's most marginalised groups. Meanwhile, the causative organism is now the world's leading infectious killer and dramatic reductions in burden will be necessary if the bold new End TB Targets are to be realised.[1] Heterogeneity in disease distribution increases as the burden of an infectious disease declines and becomes more unevenly distributed across space or social networks[2] – a phenomenon which is well recognised in the case of diseases such as malaria.[3] There are many reasons to suspect that TB epidemics are highly heterogeneous, such as the prominence of highly localised or household transmission, the wide geographical variation in disease burden within and between countries and the many individual-level factors strongly associated with risk of disease. Here we describe key drivers of heterogeneity in TB burden, discuss the challenges in quantifying this heterogeneity and consider implications for transmission dynamics and the design of interventions.

Drivers of Heterogeneity

Risk of infectious disease is dependent on characteristics of the infectious host, the organism, the susceptible host and the environment (Figure 1, Table). The complex interplay between the pathogen and the host's immune system and the propensity for *Mycobacterium tuberculosis* (*Mtb)* to enter a latent state following infection mean that many exposed individuals will never progress to active TB disease. Therefore, individual characteristics that predispose to susceptibility to infection, progression to disease after infection and infectiousness during disease episodes all contribute to heterogeneity, although the risk factors associated with each differ considerably. For example, risk of exposure is driven by sociodemographic factors (e.g. crowding, contact patterns), susceptibility to infection once exposed is influenced by processes that impair local immune responses (e.g. smoking), progression to disease may reflect systemic immune status (e.g. HIV, nutrition) and likelihood of onward transmission may be altered by cough symptomatology and disease duration (e.g. through access to care).

The Infectious Host

Medical and demographic factors also strongly influence the extent to which each affected person propagates *Mtb* infection. Smear-positive adults and particularly those with cavitary pulmonary tuberculosis transmit infection more extensively,[4] while many others, such as those with only extrapulmonary involvement, may infect no-one. Although children and persons with HIV are less likely to transmit, the degree of infectiousness is variable, with children aged over ten more often manifesting adult forms of TB.[5,6] Despite its limitations, smear microscopy remains the mainstay of TB diagnosis worldwide with advantages that include its ability to identify highly infectious individuals. Social factors such as mixing patterns also influence spread by modifying the number of contacts exposed and these patterns also differ by setting (e.g. household, workplace, general community). Importantly, social mixing patterns may act differently for *Mtb* than for other infections, given that *Mtb*, unlike many other major pathogens, is airborne and so can be transmitted without the need for direct person-to-person contact. However, the rate of transmission per day infectious is considerably lower than for other respiratory pathogens (e.g. measles, influenza),[7] meaning that amplifying factors such as cough characteristics, ability to generate

aerosols of appropriate size[8] and environmental factors may strongly influence whether infection occurs. Finally, myriad programmatic and social factors delay diagnosis and so prolong the infectious period and increase the duration of exposure,[9] thereby potentiating heterogeneity through their impact on the most marginalised groups.

The Infecting Organism

Mtb is a clonal pathogen that displays variable fitness and a complex interaction with its human host.[10] Its multiple lineages differ in their genomic make-up and in several aspects of their clinical and epidemiological behaviour, including disease progression, disease severity, transmissibility and geographic distribution (Supplemental bibliography). With recent advances in molecular epidemiology, the influence of *Mtb* genetic diversity on the outcomes of TB infection and disease is increasingly recognised. Strains are thought to have adapted to the human population they affect,[11] resulting in a sympatric relationship whereby co-evolved host populations show high rates of TB due to certain strains, but concentration within high-risk groups elsewhere.[12] However, the discordance in findings between settings and the complex interaction between pathogen, host and environment remain challenges to understanding these processes.

Arguably, the most critical form of pathogen-related heterogeneity is drug resistance, which makes clinical management considerably more challenging and expensive. Epidemiologically, transmission cycles of drug-resistant TB (DR-TB) differ from those of drug-susceptible TB because of limited access to the diagnostics available for determining drug resistance, the long duration of DR-TB treatment and clustering of DR-TB patients in high-risk settings. All these factors may act to prolong the infectious period, sustaining transmission chains of DR-TB.[13] Resistance-conferring mutations may be offset by associated physiological impairments in the organism which limit its ability to survive and reproduce (“fitness costs”), although sustained drug exposure may select for bacteria with compensatory mutations.[14] Moreover, fitness costs are likely to vary according to the drug in question (e.g. higher for rifampicin resistance than for isoniazid or streptomycin),[15] while both modelling studies and large-scale outbreaks highlight the potential for DR-TB to proliferate.[16]

The Susceptible Host

Characteristics of the susceptible host also markedly influence the likelihood of disease following exposure, which may reflect both susceptibility to infection or greater risk of progression to disease for those infected. Patterns of reactivation differ markedly by age, and comorbidities such as HIV, diabetes, malnutrition and heavy alcohol are critical considerations in the variation of risk of disease progression observed (Supplemental bibliography). For example, HIV is the strongest individual-level risk factor and a major driver of the TB epidemic in many parts of Africa, while the rising global prevalence of non-communicable diseases (e.g. diabetes) may hinder our ability to achieve control targets by impairing host immunity at the population level.[17] History of exposure and disease are also important, as people who are latently infected likely have partial protection against reinfection with the pathogen,[18] whereas previously treated persons are likely to be at substantially increased risk for recurrent disease.[19] This latter increase in risk may reflect repeated exposure, incomplete treatment, or underlying immunological vulnerability.[20]

The Physical Environment

The setting in which TB is transmitted is also an important modifier of spread – either due to increased population density, congregation of individuals with higher rates of specific risk factors or directly through environmental features that facilitate airborne transmission. Characteristics of the physical environment that may contribute to transmission include crowding, poor ventilation and high levels of indoor air pollution.[21] Furthermore, locations with these characteristics (e.g. clinics, public transit, churches, prisons, mines and informal drinking spaces) are often frequented by the same high-risk individuals, further fuelling heterogeneous transmission in these sites. These locations are themselves likely to be in close proximity, enhancing transmission in impoverished areas[22] and sustaining the epidemic.[23]

Structural and Social Determinants

Heterogeneity at the community level is driven by a complex network of proximal and distal determinants that may not always be fully explained by quantifiable risk factors. Migration, urbanisation, demographic transition and other broad global trends combined with weak and inequitable policy and planning lead to pockets of poverty, unhealthy behaviours and weak health systems in which TB thrives.[24] Social or spatial clustering of the individual-level characteristics described in the preceding sections may magnify the effect of these risk factors through transmission, as persons contact one another more if they share similar characteristics (assortative mixing). However, understanding of the effect of the various upstream determinants responsible for driving heterogeneity in TB burden is limited by the relative paucity of modelling studies in this area.[25]

Challenges in Quantifying Heterogeneity

Although substantial between- and within-country differences in burden are frequently reported, challenges exist in interpreting the differences observed between demographic, geographical or other subdivisions of the population. Our understanding of the population-level epidemiology of TB disease relies to a large extent on cases that have sought care, received a diagnosis, and been recorded through surveillance systems or local studies. The substantial proportion of cases that does not reach this stage in many settings[1] means that our estimates of heterogeneity in burden are prone to bias (Figure 2, Panels A and B). A particular consequence of relying on data from detected cases arises from the negative correlation between TB burden and access to care, which may mask heterogeneity in disease. For example, TB prevalence surveys consistently show a male predominance among adult TB cases, but this gender gap is much smaller in notifications – suggesting that men experience a higher burden but seek or access care at a lower rate than women.[26] Similar and even stronger unobserved effects – whereby mechanisms that increase risk of TB also decrease the probability of detection – may exist for features such as socio-economic status or locality. Moreover, even if bias could be eliminated from health information systems, routinely collected data are not typically disaggregated beyond broad age categories, geographic regions and drug resistance profiles, thereby limiting our ability to observe heterogeneity between smaller sub-populations without specifically designed studies.

Much less biased measures of disease burden are available from the recent increase in TB prevalence surveys. However, prevalence surveys in the general population are expensive undertakings and typically designed to yield a relative precision of 20% to 25%,[27] limiting their

ability to discern patterns among subgroups or at the district/local level. Moreover, prevalence surveys are by design cross-sectional, meaning that they cannot provide information on heterogeneity through time without additional assumptions or repeated data collection.

One important consequence of detection bias is that clusters of notifications are difficult to interpret. Apparent hotspots of TB disease may represent either true areas of intense transmission or better diagnosis (via targeted campaigns or differential access to care), such that the areas of most intense transmission may be those with the highest notification rates in some settings and the lowest in others. Travel to access care may further exaggerate this process, creating artefactual aggregations of notifications. By contrast, heterogeneity in transmission may be masked by the often substantial latent period between infection and disease onset, during which infected individuals may relocate (Figure 2, Panels C and D). This process smooths disease distribution and obscures transmission chains, while the distribution of transmission and latent infection are even harder to observe in an era when population-wide surveys of infection are no longer undertaken.

Implications for Understanding and Modelling Transmission

The impact of heterogeneity of infectiousness is influenced by characteristics of the infectious host and the organism being transmitted, and can be explored through its specific effects on the basic reproduction number, R_0.[28] While the point estimate of R_0 is often emphasised as a measure of the expected number of secondary cases caused by an average index case in an infection-naïve population, infectiousness may more appropriately be viewed as a probability distribution across a population of individuals, each with their own expected number of secondary cases. While superspreading is clearly observable in TB genomic studies,[29] saturation of close contacts – whereby contacts occur primarily among individuals who have already been infected – may increase the importance of community transmission in high-burden settings.[30]

When heterogeneity in susceptibility to TB exists, concerns regarding the assumption of a homogeneous population parallel concepts familiar in non-communicable diseases, such as cohort selection and frailty models in survival analysis. As higher-risk individuals develop incident disease,[19,31] the incidence rate of a cohort may decline simply because those who remain susceptible have a lower average risk (Figure 2, Panels E and F). This process is disabled in models that collapse risk distributions to their mean values, leading to inaccurate simulations and biased predictions. Population-level heterogeneity in susceptibility can also induce thresholds near which small epidemiological changes will cause dramatic shifts in disease burden, leading to unanticipated effects of preventive interventions[32] and faster emergence of drug-resistant strains.[33]

Any transition rate can be affected by cohort selection, as illustrated in Figure 3. Instead of the disease incidence process discussed above, consider a cohort of individuals with active TB comprised of two groups: fast and slow care seekers. As the faster care seekers leave the cohort earlier, the overall care-seeking rate will decline over time, even though it remains constant in each group. This process complicates estimation procedures and can be especially problematic in relation to rates of infection, which are proportional to the prevalence of infectious individuals and so part of a feedback loop. Moreover, epidemiological uncertainty around the most appropriate parameter values for transmission models means that multiple parameter sets may superficially replicate

observed burden,[34] which is particularly problematic for an endemic infection with a prolonged and unpredictable latency period.

Implications for Control

Targeting Risk Groups

A consequence of the heterogeneity in transmission, infection, incidence and mortality is that benefits of interventions will differ depending on the groups targeted and the distribution of the risk factors introduced above. This consideration motivates much current TB policy, with groups at higher risk of infection, disease or poor outcomes from TB episodes, such as household contacts, children, persons living with HIV, individuals with end-stage renal disease and previously treated people identified as high-priority groups for screening and treatment of latent and active TB (Supplemental bibliography). Heterogeneity in historical TB exposure is also a focus of interventions, with many low-incidence countries targeting services to foreign-born individuals,[35] given their higher LTBI prevalence and consequent risk of reactivation. However, interventions targeted at high-risk populations have not always been successful: a trial of mass screening and preventive treatment in South African miners had no impact on TB rates,[36] because of reactivation of non-cured infections and reinfection in the context of insufficient treatment and ongoing high environmental transmission risk.[37]

Synergies with non-TB interventions

Regular interactions with the health care system for the management of chronic and non-communicable diseases offer the opportunity for intensified case finding efforts, given that many such conditions increase TB risk or co-occur in populations with such increased risk. More broadly, strengthening health systems for both TB and non-communicable disease control provides the potential for synergistic interventions across diseases,[38] while improving control by addressing distal determinants should also be a high priority.[39] The observation that both historical and more recent declines[24,40] in TB burden have usually been achieved in the context of improvements in socio-economic indicators highlights the importance of such upstream determinants and is particularly relevant in the Sustainable Development Goals era.

Geographical Targeting

TB incidence shows considerable geographical clustering at multiple resolutions[41] and spatial targeting of interventions has the potential to achieve major reductions in burden through focusing on geographically discernible TB hotspots,[42] although the extent of mixing between hotspots and the broader population is important to quantify as it will modify the impact of such interventions.[43] Intensive TB control interventions targeted at Inuit communities in northern Canada, Alaska and Greenland were effective at substantially reducing the extreme rates of TB incidence and mortality observed in the 1950s.[44] New and emerging analytic tools offer opportunities to identify and quantify TB hotspots, such as a recent genomic analysis in Peru that highlighted the spatial aggregation of multidrug-resistant genotypes.[23]

Effect of Interventions on Heterogeneity

Where substantial reductions in TB burden are achieved, heterogeneity in TB distribution may increase, as transmission becomes more localised to remaining regions and population groups with fewer resources, limited healthcare access, and insufficient adherence to policy. However, even

when fully implemented, control efforts may increase or decrease transmission heterogeneity depending on the intervention design. Interventions directed at those with poor access to care and so high burden of disease may reduce heterogeneity, whereas interventions that strengthen routine programmatic management may increase heterogeneity even while decreasing overall burden. Heterogeneity may modify the impact of both targeted and untargeted interventions depending on the background burden of disease. For example, successful detection and treatment of a single active case may eliminate transmission from a community in a low-burden setting, whereas this would be harder to achieve in a high-burden setting. This may lead to unexpected relationships between control efforts and consequent reduction in the annual risk of *Mtb* infection.[45]

Economic and Equity Concerns

The targeting of TB control interventions to those with high rates of infection or disease is expected to increase the effectiveness of interventions. Consequent gains in efficiency will depend on coverage levels, accessibility, disease prevalence and contribution to transmission in the wider population of the target group. There are economies of scale to be achieved when increasing coverage, yet at high levels of coverage or for difficult to reach populations, targeted strategies may require additional supporting activities and so increase resource needs. For example, the cost-effectiveness of active case finding strategies is driven by both the heterogeneity in disease rates and in the cost of reaching different subgroups.[46] While maximising impact within a given budget is a key objective in priority setting, heterogeneity in burden, health care access and financial resources are linked to equity concerns in resource allocation for TB control strategies. Conceptually, the difference between inequalities and inequities is a value judgement about whether the observed heterogeneity is considered fair. Policy makers should seek to ensure that populations already experiencing increases in risk due to socioeconomic or other conditions (e.g., crowding, incarceration) do not experience additional disparities in access to TB diagnosis and treatment, financial burden of illness, or unwarranted exposure to infection. While the reduction of such disparities is a key policy objective, there are situations in which achieving it may imply trade-offs in efficiency gains. For example, interventions aiming to place new technologies at decentralised locations may not be as cost-effective as placement at higher levels of the health system, yet may still be prioritised to reduce social inequities in financial burden, health outcomes and access to health services.[47]

Ways Forward and Conclusions

Causes of heterogeneity in TB epidemiology are diverse and include characteristics of the infectious host, pathogen, susceptible host, environment and distal determinants – factors which may interact to amplify or reduce heterogeneity. Observed heterogeneity may not reflect reality and targeted epidemiological studies to quantify disease burden in more detail would be valuable, e.g. prevalence surveys powered to obtain precise estimates of disease burden in specific population risk groups and age-groups.

All TB modelling studies must judge which aspects of heterogeneity are sufficiently important to include given the question posed and the local context, and which should not be specifically incorporated for parsimony. This highlights the importance of: 1) detailed, context-specific data, 2) refining parameter estimation through epidemiological research, 3) communicating uncertainty in

predictive modelling and 4) confirmation of the predicted effectiveness and cost of interventions through operational research.

Heterogeneity has implications for the effectiveness and efficiency of control interventions. Targeting of interventions is an appropriate consideration in designing intervention strategies, although evidence to support specific targeted approaches is sometimes weak or contradictory. Therefore, such strategies must be considered in the context of resource availability and the ethical imperative to ensure universal access to high-quality care. Moreover, it is also important to balance the need for clear guidelines that can facilitate the broad implementation of interventions at a national or global level against the importance of developing interventions that are targeted towards specific characteristics of regional or local epidemics.

As the global TB control community looks towards ending TB, understanding and harnessing heterogeneity to improve control will become increasingly important. Key considerations in addressing heterogeneity include better assessment of disease burden in population subgroups, context-specific modelling, targeting of interventions and a focus on distal determinants of inequities in health status.

Author contributions

JMT, PJD, YAM and DWD coordinated this project. MGMG drafted the Implications for Transmission Dynamics section. GG drafted the Economic and Equity Concerns section. RMGJH drafted the Physical Environment section. NAM drafted the Effect of Interventions on Heterogeneity section. NA drafted the Targeting Risk Groups section. SS drafted the Infectious Host section. YAM drafted the Ways Forward and Conclusions section. JMT wrote the first draft of the full manuscript, which all authors reviewed and edited before submission.

Funding

No funding was received for this project. This perspective piece arose out of discussions of the Modelling Research Group at the annual meeting of the TB Modelling and Analysis Consortium (TB-MAC), which is funded by the Bill & Melinda Gates Foundation. JMT is a recipient of an Early Career Fellowship from the National Health and Medical Research Council, Australia (APP1142638). RMGJH is funded by a Starting Grant from the European Research Council (Action Number 757699). PJD is supported by the UK MRC (MR/P022081/1).

Conflicts of interest

The authors declare no conflict of interests.

References

1. World Health Organization. Global Tuberculosis Report 2017. Geneva, Switzerland: 2017. Available at: http://www.who.int/tb/publications/global_report/en/.

2. Cohen T, Colijn C, Finklea B, Murray M. Exogenous re-infection and the dynamics of tuberculosis epidemics: local effects in a network model of transmission. J R Soc Interface **2007**; 4:523–531. Available at: http://rsif.royalsocietypublishing.org/content/4/14/523.full.pdf.

3. Feachem RGA, Phillips AA, Hwang J, et al. Shrinking the malaria map: progress and prospects. Lancet (London, England) **2010**; 376:1566–78. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21035842. Accessed 15 January 2018.

4. Melsew YA, Doan TN, Gambhir M, Cheng AC, McBryde E, Trauer JM. Risk factors for infectiousness of patients with tuberculosis: a systematic review and meta-analysis. Epidemiol Infect **2018**;

5. Marais BJ, Gie RP, Schaaf HS, et al. The natural history of childhood intra-thoracic tuberculosis: a critical review of literature from the pre-chemotherapy era. Int J Tuberc Lung Dis **2004**; 8:392–402. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15141729.

6. Carvalho AC, DeRiemer K, Nunes ZB, et al. Transmission of Mycobacterium tuberculosis to contacts of HIV-infected tuberculosis patients. Am J Respir Crit Care Med **2001**; 164:2166–71. Available at: http://www.atsjournals.org/doi/abs/10.1164/ajrccm.164.12.2103078. Accessed 15 April 2018.

7. Sepkowitz KA. How Contagious Is Tuberculosis? Clin Infect Dis **1996**; 23:954–962. Available at: https://academic.oup.com/cid/article-lookup/doi/10.1093/clinids/23.5.954. Accessed 5 June 2018.

8. Turner RD, Bothamley GH. Cough and the Transmission of Tuberculosis. J Infect Dis **2015**; 211:1367–1372. Available at: https://academic.oup.com/jid/article-lookup/doi/10.1093/infdis/jiu625. Accessed 5 June 2018.

9. Lin X, Chongsuvivatwong V, Lin L, Geater A, Lijuan R. Dose–response relationship between treatment delay of smear-positive tuberculosis patients and intra-household transmission: a cross-sectional study. Trans R Soc Trop Med Hyg **2008**; 102:797–804. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18513768. Accessed 22 August 2017.

10. Coscolla M. Biological and Epidemiological Consequences of MTBC Diversity. In: Strain Variation in the Mycobacterium tuberculosis Complex: Its Role in Biology, Epidemiology and Control. Springer, 2017: 95–116.

11. Gagneux S. Host-pathogen coevolution in human tuberculosis. Philos Trans R Soc B Biol Sci **2012**; 367:850–859. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22312052. Accessed 25 March 2018.

12. Gagneux S, DeRiemer K, Van T, et al. Variable host-pathogen compatibility in Mycobacterium

tuberculosis. Proc Natl Acad Sci **2006**; 103:2869–2873. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16477032. Accessed 25 March 2018.

13. Shrestha S, Knight GM, Fofana M, et al. Drivers and trajectories of resistance to new first-line drug regimens for tuberculosis. Open forum Infect Dis **2014**; 1:ofu073. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25734143. Accessed 25 March 2018.

14. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The Competitive Cost of Antibiotic Resistance in Mycobacterium tuberculosis. Science (80-) **2006**; 312:1944–1946. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16809538. Accessed 25 March 2018.

15. Denkinger CM, Pai M, Dowdy DW. Do We Need to Detect Isoniazid Resistance in Addition to Rifampicin Resistance in Diagnostic Tests for Tuberculosis? PLoS One **2014**; 9:e84197. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24404155. Accessed 30 April 2018.

16. Knight GM, Colijn C, Shrestha S, et al. The Distribution of Fitness Costs of Resistance-Conferring Mutations Is a Key Determinant for the Future Burden of Drug-Resistant Tuberculosis: A Model-Based Analysis. Clin Infect Dis **2015**; 61:S147–S154. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26409276. Accessed 25 March 2018.

17. Odone A, Houben RMGJ, White RG, Lönnroth K. The effect of diabetes and undernutrition trends on reaching 2035 global tuberculosis targets. Lancet Diabetes Endocrinol **2014**; 2:754–764. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25194888. Accessed 4 August 2017.

18. Andrews JR, Noubary F, Walensky RP, Cerda R, Losina E, Horsburgh CR. Risk of progression to active tuberculosis following reinfection with Mycobacterium tuberculosis. Clin Infect Dis **2012**; 54:784–791. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22267721.

19. Gomes MGM, Aguas R, Lopes JS, et al. How host heterogeneity governs tuberculosis reinfection? Proc Biol Sci **2012**; 279:2473–2478. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22357260.

20. Marx FM, Floyd S, Ayles H, Godfrey-Faussett P, Beyers N, Cohen T. High burden of prevalent tuberculosis among previously treated people in Southern Africa suggests potential for targeted control interventions. Eur Respir J **2016**; 48:1227–1230. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27390274. Accessed 5 June 2018.

21. Lin H-H, Ezzati M, Murray M. Tobacco Smoke, Indoor Air Pollution and Tuberculosis: A Systematic Review and Meta-Analysis. PLoS Med **2007**; 4:e20. Available at: http://dx.plos.org/10.1371/journal.pmed.0040020. Accessed 26 April 2018.

22. Murray EJ, Marais BJ, Mans G, et al. A multidisciplinary method to map potential tuberculosis transmission 'hot spots' in high-burden communities. Int J Tuberc Lung Dis **2009**; 13:767–74. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19460255. Accessed 26 April 2018.

23. Zelner JL, Murray MB, Becerra MC, et al. Identifying Hotspots of Multidrug-Resistant Tuberculosis Transmission Using Spatial and Molecular Genetic Data. J Infect Dis **2016**; 213:287–294. Available at: https://academic.oup.com/jid/article-

lookup/doi/10.1093/infdis/jiv387. Accessed 26 March 2018.

24. Lonnroth K, Jaramillo E, Williams BG, Dye C, Raviglione M. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. Soc Sci Med **2009**; 68:2240–2246. Available at: https://www.ncbi.nlm.nih.gov/pubmed/19394122.

25. Pedrazzoli D, Boccia D, Dodd PJ, et al. Modelling the social and structural determinants of tuberculosis: opportunities and challenges. Int J Tuberc Lung Dis **2017**; 21:957–964. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28826444. Accessed 25 March 2018.

26. Horton KC, MacPherson P, Houben RMGJ, White RG, Corbett EL. Sex Differences in Tuberculosis Burden and Notifications in Low- and Middle-Income Countries: A Systematic Review and Meta-analysis. PLOS Med **2016**; 13:e1002119. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27598345. Accessed 26 March 2018.

27. World Health Organization. Tuberculosis Prevalence Surveys: a Handbook. 2011.

28. Colijn C, Cohen T, Murray M. Emergent heterogeneity in declining tuberculosis epidemics. J Theor Biol **2007**; 247:765–774. Available at: http://ac.els-cdn.com/S002251930700197X/1-s2.0-S002251930700197X-main.pdf?_tid=701971423bef0a352e3f5d824a965769&acdnat=1345012153_4a5156ce5fbcc81695d4f605ae8bcb3a.

29. Ypma RJ, Altes HK, van Soolingen D, Wallinga J, van Ballegooijen WM. A sign of superspreading in tuberculosis: highly skewed distribution of genotypic cluster sizes. Epidemiology **2013**; 24:395–400. Available at: https://www.ncbi.nlm.nih.gov/pubmed/23446314.

30. McCreesh N, White RG. An explanation for the low proportion of tuberculosis that results from transmission between household and known social contacts. Sci Rep **2018**; 8:5382. Available at: http://www.nature.com/articles/s41598-018-23797-2. Accessed 8 June 2018.

31. Gomes MGM, Barreto ML, Glaziou P, et al. End TB strategy: the need to reduce risk inequalities. BMC Infect Dis **2016**; 16:132. Available at: http://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-016-1464-8. Accessed 22 May 2018.

32. Gomes MGM, Rodrigues P, Hilker FM, et al. Implications of partial immunity on the prospects for tuberculosis control by post-exposure interventions. J Theor Biol **2007**; 248:608–617. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17669435. Accessed 27 March 2018.

33. Cohen T, Lipsitch M, Walensky RP, Murray M. Beneficial and perverse effects of isoniazid preventive therapy for latent tuberculosis infection in HIV-tuberculosis coinfected populations. Proc Natl Acad Sci U S A **2006**; 103:7042–7047. Available at: http://www.pnas.org/content/103/18/7042.full.pdf.

34. Ragonnet R, Trauer JM, Scott N, Meehan MT, Denholm JT, McBryde ES. Optimally capturing latency dynamics in models of tuberculosis transmission. Epidemics **2017**; 21:39–47.

Available at: http://www.sciencedirect.com/science/article/pii/S1755436517300178. Accessed 25 June 2017.

35. Centers for Disease Control and Prevention. Reported Tuberculosis in the United States. 2016. Available at: https://www.cdc.gov/tb/statistics/reports/2016/default.htm. Accessed 26 March 2018.

36. Churchyard GJ, Fielding KL, Lewis JJ, et al. A Trial of Mass Isoniazid Preventive Therapy for Tuberculosis Control. N Engl J Med **2014**; 370:301–310. Available at: http://www.nejm.org/doi/10.1056/NEJMoa1214289. Accessed 30 April 2018.

37. Vynnycky E, Sumner T, Fielding KL, et al. Tuberculosis Control in South African Gold Mines: Mathematical Modeling of a Trial of Community-Wide Isoniazid Preventive Therapy. Am J Epidemiol **2015**; 181:619–632. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25792607. Accessed 2 October 2018.

38. Pan S-C, Ku C-C, Kao D, Ezzati M, Fang C-T, Lin H-H. Effect of diabetes on tuberculosis control in 13 countries with high tuberculosis: a modelling study. lancet Diabetes Endocrinol **2015**; 3:323–30. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25754415. Accessed 26 March 2018.

39. Carter DJ, Glaziou P, Lönnroth K, et al. The impact of social protection and poverty elimination on global tuberculosis incidence: a statistical modelling analysis of Sustainable Development Goal 1. Lancet Glob Heal **2018**; 6:e514–e522. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29580761. Accessed 6 June 2018.

40. Lienhardt C, Glaziou P, Uplekar M, Lönnroth K, Getahun H, Raviglione M. Global tuberculosis control: lessons learnt and future prospects. Nat Rev Microbiol **2012**; 10:407–416. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22580364. Accessed 3 October 2018.

41. Shaweno D, Karmakar M, Alene KA, et al. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. BMC Med **2018**;

42. Dowdy DW, Golub JE, Chaisson RE, Saraceni V. Heterogeneity in tuberculosis transmission and the role of geographic hotspots in propagating epidemics. Proc Natl Acad Sci U S A **2012**; Available at: http://www.ncbi.nlm.nih.gov/pubmed/22645356.

43. Shaweno D, Trauer JM, Denholm JT, McBryde ES. A novel Bayesian geospatial method for estimating tuberculosis incidence reveals many missed TB cases in Ethiopia. BMC Infect Dis **2017**; 17.

44. Grzybowski S, Dorken E. Tuberculosis in Inuit. Ecol Dis **1983**; 2:145–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/6332010. Accessed 26 March 2018.

45. Uys P, Marais BJ, Johnstone-Robertson S, Hargrove J, Wood R. Transmission Elasticity in Communities Hyperendemic for Tuberculosis. Clin Infect Dis **2011**; 52:1399–1404. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21628479. Accessed 26 April 2018.

46. Azman AS, Golub JE, Dowdy DW. How much is tuberculosis screening worth? Estimating the value of active case finding for tuberculosis in South Africa, China, and India. BMC Med **2014**; 12:216. Available at: http://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-014-0216-0. Accessed 26 March 2018.

47. Sachdeva KS, Raizada N, Sreenivas A, et al. Use of Xpert MTB/RIF in Decentralized Public Health Settings and Its Effect on Pulmonary TB and DR-TB Case Finding in India. PLoS One **2015**; 10:e0126065. Available at: http://dx.plos.org/10.1371/journal.pone.0126065. Accessed 26 March 2018.

48. Gardy JL, Johnston JC, Sui SJH, et al. Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak. N Engl J Med **2011**; 364:730–739. Available at: http://www.nejm.org/doi/abs/10.1056/NEJMoa1003176. Accessed 27 September 2018.

49. Valway SE, Sanchez MPC, Shinnick TF, et al. An Outbreak Involving Extensive Transmission of a Virulent Strain of Mycobacterium tuberculosis. N Engl J Med **1998**; 338:633–639. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9486991. Accessed 27 September 2018.

50. Gandhi NR, Weissman D, Moodley P, et al. Nosocomial Transmission of Extensively Drug-Resistant Tuberculosis in a Rural Hospital in South Africa. J Infect Dis **2013**; 207:9–17. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23166374. Accessed 3 October 2018.

51. Verver S, Warren RM, Beyers N, et al. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. Am J Respir Crit Care Med **2005**; 171:1430–1435. Available at: http://www.ncbi.nlm.nih.gov/pubmed/15831840.

52. Nolan CM, Elarth AM, Barr H, Saeed AM, Risser DR. An Outbreak of Tuberculosis in a Shelter for Homeless Men: A Description of Its Evolution and Control. Am Rev Respir Dis **1991**; 143:257–261. Available at: http://www.ncbi.nlm.nih.gov/pubmed/1990937. Accessed 3 October 2018.

53. Jones TF, Craig AS, Valway SE, Woodley CL, Schaffner W. Transmission of tuberculosis in a jail. Ann Intern Med **1999**; 131:557–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10523215. Accessed 27 September 2018.

54. Ehrenkranz NJ, Kicklighter JL. Tuberculosis Outbreak in a General Hospital: Evidence for Airborne Spread of Infection. Ann Intern Med **1972**; 77:377. Available at: http://annals.org/article.aspx?doi=10.7326/0003-4819-77-3-377. Accessed 27 September 2018.

55. Nery JS, Rodrigues LC, Rasella D, et al. Effect of Brazil's conditional cash transfer programme on tuberculosis incidence. Int J Tuberc Lung Dis **2017**; 21:790–796. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28633704. Accessed 27 September 2018.

Figure 1. Conceptual framework for understanding heterogeneity in TB epidemiology

The cone indicates that the most local drivers are positioned towards the top of the figure and the broadest drivers towards bottom, rather than reflecting the importance of these factors.

Figure 2. Illustration of some selected concepts from the text

Panel A illustrates the degree of heterogeneity that might be observed among individuals with good access to the healthcare system (unblurred discs) compared to those with poor access (blurred discs). This may be substantially less than the heterogeneity that exists in the population as a whole (Panel B).

Panel C represents a series of transmission events and Panel D illustrates the subsequent relocation of infected and uninfected individuals. This results in a more homogeneous distribution of infection across the population at this later time point, even though transmission was highly heterogeneous.

Panel E represents a series of individuals at variable risk of infection and Panel F illustrates selection of higher risk individuals through the infection process. Although infection is the selecting illustrated process here, similar principles would apply to progression from infection to disease, through stages of the disease process and to interaction with the health system.

Figure 3. Composition of a simple two-stratum heterogeneous cohort over time from entry to an epidemiological state (active undiagnosed TB)

Plot displays the percentage of patients with active tuberculosis remaining undiagnosed after the onset of infectiousness (time 0 on the horizontal axis), under the assumption that 50% of the initial cohort has an average duration of infectiousness of one month ("high rate group", blue), and 50% of the cohort has a duration of infectiousness of 6 months ("low rate group", red). The true total percentage of patients remaining infectious with time since onset of infectiousness (solid line) is compared against: the proportion that would be expected to remain if the whole cohort was assumed to have the average time to diagnosis (3.5 months); and the proportion that would be expected to remain if the whole cohort was assumed to have a rate of diagnosis that is the average of the rates of the two groups (dotted line). The amount of the total population comprised of high-rate and low-rate persons at each time point is indicated by coloured shading, demonstrating that the remaining cohort is increasingly comprised of low-rate individuals over time.

Main messages

- Although often masked by reported aggregate estimates, the distribution of TB is heterogeneous and harnessing this heterogeneity may be critical to further progress in the fight against TB.
- Drivers of heterogeneity in TB burden include characteristics of the organism, infectious host, susceptible host, environment and distal determinants.
- More detailed epidemiological data are needed to define and quantify this variation.
- Quantification of heterogeneity in TB distribution is complicated by heterogeneity in the process of detecting cases.
- Incorporating heterogeneity in TB transmission models is necessary when capturing epidemiological phenomena that include superspreading and cohort selection.
- Targeting high-risk groups is an established approach and is an important consideration in designing control interventions, but may not always improve effectiveness and may incur additional costs. However, targeting interventions should be considered in the context of ethical and equity concerns, programmatic efficiency and synergies across the broader health system.

Table. Examples of Specific Forms of Heterogeneity and Ways Forward

Source of Heterogeneity	Examples of existing evidence	Data needs	Analytic needs	Intervention needs
The Infectious Host	Sequencing and social network analysis suggest that some individuals may act as "superspreaders"[48]	Importance of biological variables, e.g. aerosolisation, cough frequency	Implications of hosts with differential infectiousness and superspreading	Tools to identify the most infectious patients
	Available data on contact patterns (principally from low-burden settings) suggest age-specific (assortative) mixing	Data on contact patterns from high-burden settings and for risk factors relevant to TB (e.g. HIV status)	Importance of population groups to sustaining transmission relative to their burden of disease	Case-finding efforts designed to identify patients with high-risk mixing patterns for broader dissemination of infection
The Infecting Organism	Strain responsible for extensive community spread confirmed to be highly virulent in mouse model[49]	Mechanisms of strain diversity and virulence	Implications of selecting for strains of greater fitness	Interventions to limit infectiousness of difficult-to-treat strains
	Highly resistant forms of TB causing extensive outbreaks, e.g. XDR-TB in Tugela Ferry, South Africa[50]	Fitness costs associated with drug resistance	Likely future trajectory of drug resistance	Improved identification and treatment of highly transmissible strains of drug-resistant tuberculosis
The Susceptible Host	Individuals previously treated for TB had higher rates of recurrent TB due to reinfection than the general population in Cape Town, South Africa[51]	Protection or susceptibility afforded by past TB episodes and whether this is attributable to infection or progression risk	Distinguish the individual-level effect of increased susceptibility post-disease episode from the effect of selecting for a more susceptible cohort through infection	Protection of highest risk individuals from infection or progression to disease

	Specific risk groups may experience polyclonal outbreaks[52]	Better estimates of disease prevalence in risk groups	Anticipated effects of trends in comorbid risk factors on TB	TB control interventions that link with systems for other high-risk conditions
The Physical Environment	Incarceration may have been a significant driver of community transmission [53]	Better estimates of location-specific TB transmission risk	Valid models for translating environmental heterogeneity into transmission risk	Active case-finding targeted at high-risk environments (e.g. prisons, transit)
	Greater proportion of infected contacts in less well ventilated hospital wards[54]	Ability of specific interventions (e.g. improved ventilation) to reduce that risk	Projected population-level impact of targeted environmental interventions	Mitigation of TB transmission through modification of high-risk built environments
Distal Determinants	Ecological observation of declining TB rates during times of improvements in living standards[40]	Mechanistic linkages between poverty alleviation and TB transmission	Projected ability of social protection and similar efforts to reduce heterogeneity	Linkage between TB control programs and schemes to alleviate poverty and/or address other distal determinants
	Association between coverage of Brazil's conditional cash transfer program and improved TB control[55]	TB-specific effects of broader interventions	Models of the impact of TB on other outcomes in vulnerable populations	Implementation of TB interventions in a fashion that mitigates burden on the highest risk populations, thus promoting equity and reducing disparities in risk

XDR-TB, extensively drug-resistant tuberculosis.

Figure 1

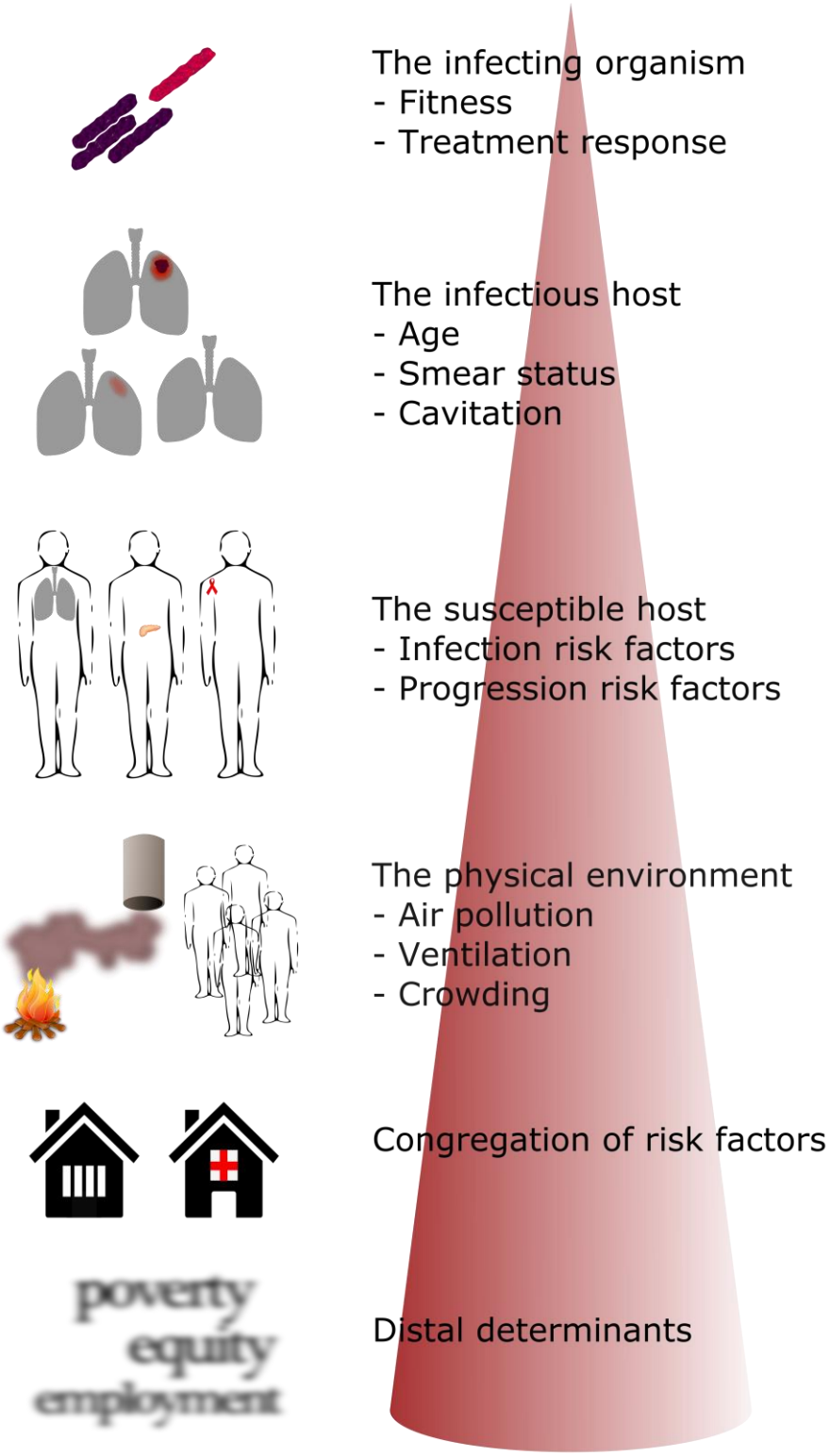

Figure 2

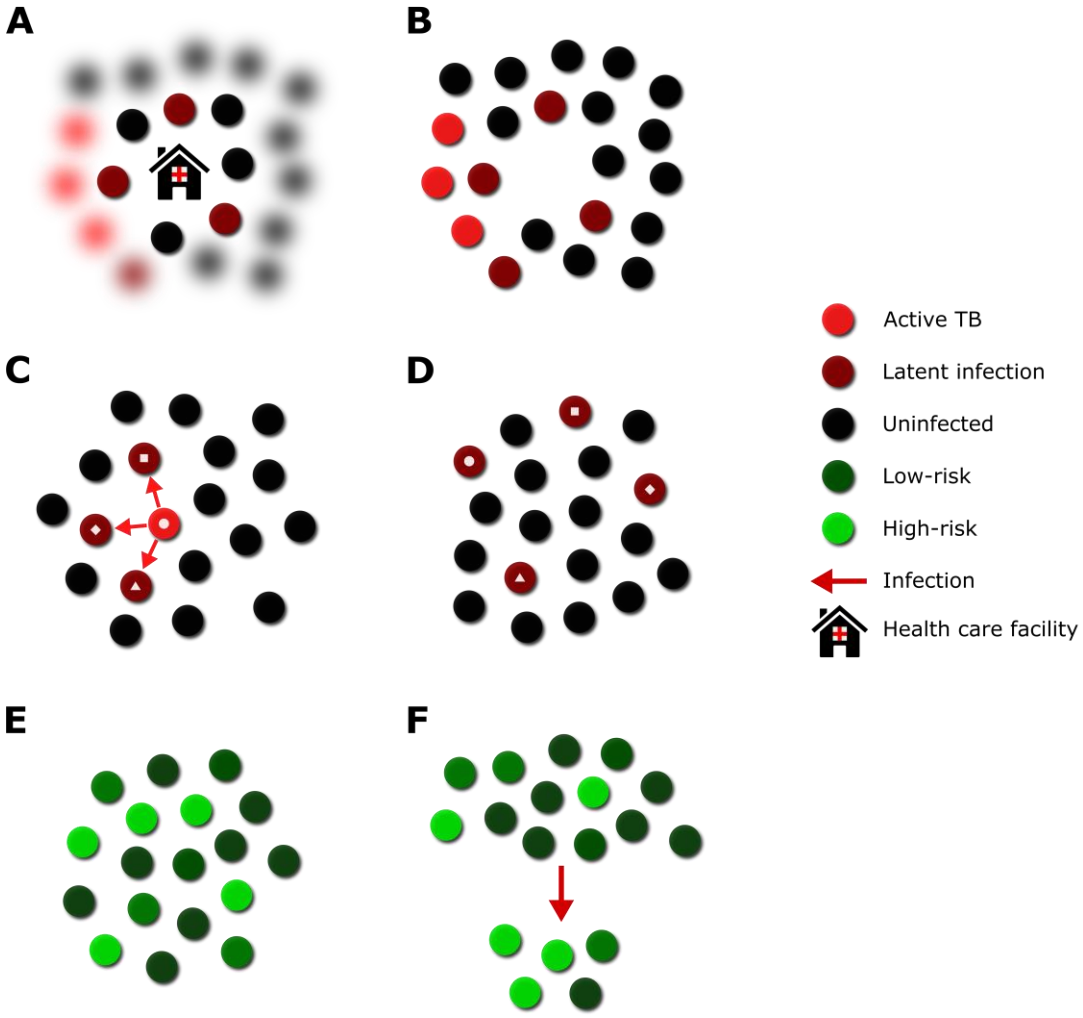

Figure 3